VAIDS-DSML exercise block 10

Jasmin Duvivié

FH TECHNIKUM WIEN
University of Applied Sciences

k-means clustering

- problem

- 660 samples of vertebral columns

- 7 features

- no target

- we need to group the samples into meaningful clusters

- solution

- visualize the features to check which is suitable for clustering

- set random values as cluster centers

- assign samples to clusters based on minimal Euclidian distance

- repeat until clusters don't significantly improve

k-means clustering 1ˢᵗ iteration